

A First Course in Programming and Statistics

The Art of Data Science

Due to its data handling and modeling capabilities as well as its flexibility, R is becoming the most widely used software in bioinformatics. R Programming for Bioinformatics explores the programming skills needed to use this software tool for the solution of bioinformatics and computational biology problems. Drawing on the author's first-hand experiences as an expert in R, the book begins with coverage on the general properties of the R language, several unique programming aspects of R, and object-oriented programming in R. It presents methods for data input and output as well as database interactions. The author also examines different facets of string handling and manipulations, discusses the interfacing of R with other languages, and describes how to write software packages. He concludes with a discussion on the debugging and profiling of R code. With numerous examples and exercises, this practical guide focuses on developing R programming skills in order to tackle problems encountered in bioinformatics and computational biology.

Discover best practices for data analysis and software development in R and start on the path to becoming a fully-fledged data scientist. This book teaches you techniques for both data manipulation and visualization and shows you the best way for developing new software packages for R. Beginning Data Science in R details how data science is a combination of statistics, computational science, and machine learning. You'll see how to efficiently structure and mine data to extract useful patterns and build mathematical models. This requires computational methods and programming, and R is an ideal programming language for this. This book is based on a number of lecture notes for classes the author has taught on data science and statistical programming using the R programming language. Modern data analysis requires computational skills and usually a minimum of programming. What You Will Learn Perform data science and analytics using statistics and the R programming language Visualize and explore data, including working with large data sets found in big data Build an R package Test and check your code Practice version control Profile and optimize your code Who This Book Is For Those with some data science or analytics background, but not necessarily experience with the R programming language.

The Art of R ProgrammingA Tour of Statistical Software DesignNo Starch Press

The problem of inducing, learning or inferring grammars has been studied for decades, but only in recent years has grammatical inference emerged as an independent field with connections to many scientific disciplines, including bio-informatics, computational linguistics and pattern recognition. This book meets the need for a comprehensive and unified summary of the basic techniques and results, suitable for researchers working in these various areas. In Part I, the objects of use for grammatical inference are studied in detail: strings and their topology, automata and grammars, whether probabilistic or not. Part II carefully explores the main questions in the field: What does learning mean? How can we associate complexity theory with learning? In Part III the author describes a number of techniques and algorithms that allow us to learn from text, from an informant, or through interaction with the environment. These concern automata, grammars, rewriting systems, pattern languages or transducers.

A Step-by-Step Function Guide to Data Analysis

The Age of Spiritual Machines

Proven Recipes for Data Analysis, Statistics, and Graphics

The Art of R Programming

Using R for Introductory Statistics

Computational statistics and statistical computing are two areas that employ computational, graphical, and numerical approaches to solve statistical problems, making the versatile R language an ideal computing environment for these fields. One of the first books on these topics to feature R, Statistical Computing with R covers the traditiona This book provides an elementary-level introduction to R, targeting both non-statistician scientists in various fields and students of statistics. The main mode of presentation is via code examples with liberal commenting of the code and the output, from the computational as well as the statistical viewpoint. Brief sections introduce the statistical methods before they are used. A supplementary R package can be downloaded and contains the data sets. All examples are directly runnable and all graphics in the text are generated from the examples. The statistical methodology covered includes statistical standard distributions, one- and two-sample tests with continuous data, regression analysis, one-and two-way analysis of variance, regression analysis, analysis of tabular data, and sample size calculations. In addition, the last four chapters contain introductions to multiple linear regression analysis, linear models in general, logistic regression, and survival analysis.

The high-level language of R is recognized as one of the most powerful and flexible statistical software environments, and is rapidly becoming the standard setting for quantitative analysis, statistics and graphics. R provides free access to unrivalled coverage and cutting-edge applications, enabling the user to apply numerous statistical methods ranging from simple regression to time series or multivariate analysis. Building on the success of the author's bestselling Statistics: An Introduction using R, The R Book is packed with worked examples, providing an all inclusive guide to R, ideal for novice and more accomplished users alike. The book assumes no background in statistics or computing and introduces the advantages of the R environment, detailing its applications in a wide range of disciplines. Provides the first comprehensive reference manual for the R language, including practical guidance and full coverage of the graphics facilities. Introduces all the statistical models covered by R, beginning with simple classical tests such as chi-square and t-test. Proceeds to examine more advance methods, from regression and analysis of variance, through to generalized linear models, generalized mixed models, time series, spatial statistics, multivariate statistics and much more. The R Book is aimed at undergraduates, postgraduates and professionals in science, engineering and medicine. It is also ideal for students and professionals in statistics, economics, geography and the social sciences.

Perform data analysis with R quickly and efficiently with more than 275 practical recipes in this expanded second edition. The R language provides everything you need to do statistical work, but its structure can be difficult to master. These task-oriented recipes make you productive with R immediately. Solutions range from basic tasks to input and output, general statistics, graphics, and linear regression. Each recipe addresses a specific problem and includes a discussion that explains the solution and provides insight into how it works. If you're a beginner, R Cookbook will help get you started. If you're an intermediate user, this book will jog your memory and expand your horizons. You'll get the job done faster and learn more about R in the process. Create vectors, handle variables, and perform basic functions Simplify data input and output Tackle data structures such as matrices, lists, factors, and data frames Work with probability, probability distributions, and random variables Calculate statistics and confidence intervals and perform statistical tests Create a variety of graphic displays Build statistical models with linear regressions and analysis of variance (ANOVA) Explore advanced statistical techniques, such as finding clusters in your data

Learning Automata and Grammars

The Art of Coding

Data Analysis, Visualization, and Modelling for the Data Scientist

The Art of Assembly Language, 2nd Edition

Learning R Programming

The Art of Computer Programming, Volume 4A

In just 24 lessons of one hour or less, Sams Teach Yourself R in 24 Hours helps you learn all the R skills you need to solve a wide spectrum of real-world data analysis problems. You'll master the entire data analysis workflow, learning to build code that's efficient, reproducible, and suitable for sharing with others. This book's straightforward, step-by-step approach teaches you how to import, manipulate, summarize, model, and plot data with R; formalize your analytical code; and build powerful R packages using current best practices. Practical, hands-on examples show you how to apply what you learn. Quizzes and exercises help you test your knowledge and stretch your skills. Learn How To Install, configure, and explore the R environment, including RStudio Use basic R syntax, objects, and packages Create and manage data structures, including vectors, matrices, and arrays Understand lists and data frames Work with dates, times, and factors Use common R functions, and learn to write your own Import and export data and connect to databases and spreadsheets Use the popular tidy, dplyr and data.table packages Write more efficient R code with profiling, vectorization, and initialization Plot data and extend your graphical capabilities with ggplot2 and Lattice graphics Develop common types of models Construct high-quality packages, both simple and complex Write R classes: S3, S4, and Reference Classes Use R to generate dynamic reports Build web applications with Shiny Register your book at informit.com/register for convenient access to updates and corrections as they become available. This book's source code can be found at <http://www.mango-solutions.com/wp/teach-yourself-r-in-24-hours-book>.

Learn to expertly apply a range of machine learning methods to real data with this practical guide. Machine learning without advanced math! This book presents a serious, practical look at machine learning, preparing you for valuable insights on your own data. The Art of Machine Learning is packed with real dataset examples and sophisticated advice on how to make full use of powerful machine learning methods. Readers will need only an intuitive grasp of charts, graphs, and the slope of a line, as well as familiarity with the R programming language. You'll become skilled in a range of machine learning methods, starting with the simple k-Nearest Neighbors method (k-NN), then on to random forests, gradient boosting, linear/logistic models, support vector machines, the LASSO, and neural networks. Final chapters introduce text and image classification, as well as time series. You'll learn not only how to use machine learning methods, but also why these methods work, providing the strong foundational background you'll need in practice. Additional features:

- How to avoid common problems, such as dealing with "dirty" data and factor variables with large numbers of levels
- A look at typical misconceptions, such as dealing with unbalanced data
- Exploration of the famous Bias-Variance Tradeoff, central to machine learning, and how it plays out in practice for each machine learning method
- Dozens of illustrative examples involving real datasets of varying size and field of application
- Standard R packages are used throughout, with a simple wrapper interface to provide convenient access. After finishing this book, you will be well equipped to start applying machine learning techniques to your own datasets.

Efficient R Programming

Algorithms+Data+R